Final Technical Report

NEHRP/USGS Award Number: G17AP00005

Assessing Crustal Deformation in the San Francisco Bay Area:
Collaborative Research with the U.S. Geological Survey, Menlo Park

Mark H. Murray, Assoc. Research Professor
Department of Earth and Environmental Sciences
New Mexico Institute of Mining and Techonlogy
801 Leroy Place
Socorro, New Mexico 87801
Tel: 575-835-6930, Fax: 575-835-6436
Email: mark.murray@nmt.edu

Term covered by this award: 12/01/2016-11/30/2017

Abstract

The goal of this project is to assess crustal deformation in the San Francisco Bay area to improve constraints on fault locking and deep slip rates for seismic hazard analysis. The San Francisco Bay area is underlain by multiple, sub-parallel strands of the San Andreas fault system, such as the San Andreas, Hayward-Rodgers Creek, and Calaveras faults that accommodate much of the Pacific and North America relative plate motion. We use available geodetic data (primarily GPS) to estimate models of strain accumulation across the San Andreas fault system, and to critically assess their bounds to improve forecasts of seismic hazard in the Bay area and Delta regions. As part of the EarthScope PBO/GAGE project, we processed most of the continuously operating GNSS stations throughout the western U.S. using network methods with the GAMIT/GLOBK software. The data were also processed by using precise point positioning (PPP) methods implemented in the JPL GIPSY software. The two software packages apply different analysis strategies that yield similar position estimates but have significant differences between in their uncertainty estimates. Unlike the network method, the PPP approach is done on an individual station basis that provides no estimate of inter-station correlations due, in part, to orbital and atmospheric effects. Since strain accumulation models are primarily sensitive to the differences between station velocities, ignoring inter-station correlations may significantly bias estimated uncertainties of the strainrates and consequently the bounds on fault slip and locking. We examine the different approaches to assess the effect they have on strain accumulation modeling.

We have used the results of the PBO/GAGE PPP, network, and combined PBO position solutions to perform an initial assessement of inter-station correlations using the a one-dimensional screw dislocation model applied to a linear array across the creeping section of the San Andreas fault in central California. This simple linear inversion model for fault slip allows for direct evaluation of the effect of the different correlation matrices. Our preliminary results show differences in the formal uncertainties of the deep sliprates, but when these uncertainties are scaled to make them consistent with the residual scatter, the differences are not significant, with all solutions yielding a deep sliprate estimate of 31.1 ± 0.4 mm/yr, consistent with estimates based on more realistic finite-fault modeling studies. The estimated inter-station correlations of the fault-parallel velocity components are less than 0.25 and greater than -0.15, indicating a lack of strong correlations or anti-correlations between the stations. The network-derived correlations have slightly more positive inter-correlations (0.05-0.2) than the PPP-derived solutions. Appliction of the network-derived correlation combined with the formal uncertainties of the PPP-derived solution results in a modestly higher (10%) sliprate uncertainty. More realistic three-dimensional fault modeling may provide additional insights into whether this increase is a significant, but our preliminary conclusion is that the differences between network and PPP-derived solutions caused by inter-station correlations have a neglible effect on strain accumulation modeling estimates.

Project Description

The goal of this project is to assess crustal deformation in the San Francisco Bay area to improve constraints on fault locking and deep slip rates for seismic hazard analysis. The San Francisco Bay area is underlain by multiple, sub-parallel strands of the San Andreas fault system, such as the San Andreas, Hayward-Rodgers Creek, and Calaveras faults that accommodate much of the Pacific and North America relative plate motion. These faults are thought to pose the greatest seismic hazards in the densely populated Bay area, and consequently have been the focus of much study. Other faults, particularly in the east Bay region such as the Concord-Green Valley, Greenville, Tracy, Midland, and faults within the San Joaquin-Sacramento Delta region, are less well characterized, but potentially represent significant sources of seismic hazard, as demonstrated by the 2014 M6 earthquake near Napa. We use available geodetic data (primarily GPS) to estimate models of strain accumulation across the San Andreas fault system, and to critically assess their bounds to improve forecasts of seismic hazard in the Bay area and Delta regions.

As part of the EarthScope PBO/GAGE project, we operated an analysis center at the New Mexico Institute of Mining and Technology (NMT) that processed most of the continuously operating stations throughout the western U.S. using network methods with the GAMIT/GLOBK software. The data were also processed by Central Washington University (CWU) using precise point positioning (PPP) methods implemented in the JPL GIPSY software, which is the same approach employed by USGS Menlo Park to analyze continuous and survey-mode data collected in the western U.S. since the early 1990s. The results of NMT and CWU analysis were combined to produce time series and station velocities by an analysis center coordinator at the Massachusetts Institute of Technology. The two software packages apply different analysis strategies that yield similar position estimates but have significant differences between in their uncertainty estimates. In particular, the PPP approach is done on an individual station basis that provides no estimate of expected inter-station. Since strain accumulation models are most sensitive to the differences between station velocities (i.e., the strainrates), ignoring inter-station correlations may significantly bias estimated uncertainties of the strainrates and consequently the bounds on fault slip and locking. We examine the different approaches, and assess the effect they have on strain accumulation modeling.

During the project period, we completed analysis of over 2000 GNSS stations in the western U.S. for the PBO/GAGE project, culminating with an official release of NMT, CWU, and combined PBO daily station positions, time series, and velocity products in June 2019. We have used the results of these analyses to perform an initial assessement of inter-station correlations using the a one-dimensional screw dislocation model applied to a linear array across the creeping section of the San Andreas fault in central California. This simple linear inversion model for fault slip allows for direct evaluation of the effect of the different correlation matrices. Our preliminary results suggest that the differences do not significantly affect realistically scaled uncertainty on the estimated deep-slip rate.

In particular, this project addresses the following components of the USGS Earthquakes Hazard Program for the Northern California region. ELEMENT I. Use crustal deformation measurements to constrain regional deformation rate, fault slip rates, fault creep, fault mechanics, strain transients, and models of stress evolution for Northern California. In particular, it addresses priorities to: Determine bounds on deformation rates across the San Andreas, Hayward-Rodgers Creek, West Napa, and Calaveras-Green Valley-Bartlett Springs fault systems, and to integrate the range of geodetic observations with other available information to develop and test detailed models of crustal deformation with particular concern for the eastern Bay Area. We anticipate that it will contribute to seismic hazard assessment in the San Joaquin-Sacramento Delta region and for the Sacramento River Delta levee system by improving estimates of earthquake recurrence of active faults proximal to the San Joaquin-Sacramento Delta, in particular, the Concord-Green Valley-Bartlett Springs fault system, and the Greenville, Midland, and Tracy faults. ELEMENT III. Develop, evolve, and test probabilistic models for earthquake rupture in Northern California, in coordination with the Uniform California Earthquake Rupture Forecast (UCERF) by refining strain rates and deformation models used in assigning slip rates to faults and evaluating off-fault deformation.

Fault slip rate estimation

More than 7.5 million people live in nine counties in the San Francisco Bay area. This area is underlain by multiple strands of the ~100 km wide San Andreas fault system (SAFS), including the major San Andreas, Hayward-Rodgers Creek, and Calaveras faults, as well as less well characterized faults in the east Bay and San Joaquin-Sacramento Delta regions such as the Concord-Green Valley, Greenville, Tracy, and Midland faults (Figure 1). The Bay area has been severely damaged by several major earthquakes, including the 1868 M7 Hayward, 1906 M7.8 San Francisco, and 1989 M6.9 Loma Prieta, and significantly impacted by many smaller events such as the recent 2014 M6.0 event near Napa. The third Uniform California Earthquake Rupture Forecast (UCERF3) model estimates a 72% probability of an M6.7 event occurring in the Bay area in the next 30 years (*Field et al.,* 2015). With a dense population living in close proximity to numerous earthquake-prone faults, the San Francisco Bay area is one of the most seismically hazardous regions in the U.S., and has been the focus of intensive research to better characterize those hazards.

Geodetic measurements, which are particularly useful for monitoring deformation and strain on deep structures throughout the seismic cycle, have been conducted in the Bay area for over a 150 years since triangulation networks were established for coastal navigation. Examining displacements measured by these networks due to the 1906 San Francisco earthquake led Henry Fielding Reid to propose his elastic rebound theory. Since the early 1990s, observations using the Global Positioning System (GPS) and interferometric sythetic aperature radar (InSAR), combined with surface fault creep measurements and other observations, such as repeating earthquakes that provide constraints on fault slip at depth, have provided insights into the nature of strain accumulation on the Bay area faults, including improved estimates of deep slip rates and regions of fault locking (e.g., *d'Alessio et al.,* 2005; *Evans et al.,* 2012), postseismic relaxation following the 1989 Loma Prieta earthquake (*Burgmann et al., 1997*), and the possible linking of the Hayward and Calaveras fault (*Chaussard et al.,* 2015. Geodetic estimates of strain accumulation and off-fault strain-rates, which are now being incorporated into UCERF models, can provide important constraints for assessing the hazards posed by future earthquakes.

In the San Francisco Bay area, the San Andreas fault system accommodates ~38 mm/yr of the total 48 mm/yr right-lateral Pacific-North America relative plate motion. This relative motion is partitioned primarily onto the major San Andreas, Hayward-Rodgers Creek, and Calaveras faults., and previous geodetic studies have provided constraints on the fault slip-rates. The USGS has conducted both Geodolite (trilateration) in the 1970's and 1980's, and GPS surveys in the 1990's in the vicinity of Pt Reyes. The Geodolite measurements, summarized by *Lisowski et al*. [1991], show that the cumulative fault-parallel displacement rate across the 115 km wide network at Point Reyes is 31 ± 3 mm/yr. The preferred interseismic strain accumulation model derived a broader GPS profile extending from Pt Reyes to eastern California, assumes 20.8 ± 1.9 mm/yr on the San Andreas fault, 10.3 ± 2.6 mm/yr on the Rodgers Creek fault, and 8.1 ± 2.1 mm/yr on the Green Valley fault, with total motion across the SAFS of 39.2 ± 3.8 mm/yr over a 120-km wide zone [*Prescott et al.,* 2001]. They also find a zone of compression between the Coast Ranges and the Great Valley accommodating 3.8 ± 1.5 mm/yr of shortening over a narrow zone. These results are similar to those found from continuous GPS measurements across the central Bay Area, where the total deep slip rate is 37.2 ± 0.1 mm/yr, and the narrow zone of convergence near the Great Valley shows 2.4 ± 0.4 mm/yr of shortening [*Murray and Segall,* 2001].

These earlier studies used simple fault geometries, often assuming infinitely long strike-slip faults. More recent studies have used finite-fault (i.e., rectangular or triangular patch) geometries that can better account for complex geometries of the faults, including changes in strike and dip along the faults. In our BAVU modeling effort (*d'Allesio et al.,* 2005), we used estimated velocities of over 190 GPS stations and angular velocity-block motion models following (REF) to infer rates on the major faults, as well as the San Gregorio,and several east Bay faults, including the West Napa, Green Valley, Concord, and Greenville faults. *Evans et al*. (2012) extended this modeling, using the same BAVU GPS data plus InSAR observations, to better constrain slip variations along the Hayward fault. The Hayward and Calaveras faults have been the focus of several detailed studies that take advantage of the dense spatial coverage of InSAR observations, and in some cases the deep-slip constraints provided by repeating earthquakes (*Burgmann et al.,* 2000; *Manaker et al.,* 2003; *Schmidt et al.,* 2005; *Johanson and Burgmann*, 2005). In the last 15 years, many additional GPS stations have been installed in northern California, primarily as part of the EarthScope PBO project. In the San Francisco Bay area and San Joaquin-Sacramento Delta regions, there are now data

from over 130 continuous GPS stations including several in the Central valley that can potentially provide useful constraints on poorly constrained East Bay faults, such as the Greenville, Midway-Black Butte, and Vernalis faults near Tracy, and the Kirby Hills-Vaca and Midland faults in the Delta region.

In general, the geodetic data, given sufficient far-field site locations, robustly constraint the total interseismic deformation accommodated by the entire San Andreas fault zone, but more weakly constrain the strain accumulating on the individual, sub-parallel faults. For example, in *Freymueller et al*. [1999] we report on GPS data collected in two profiles near Ukiah and Willits that span the entire northern SAFS and include sites in eastern California, estimated interseismic deep slip rates and fault locking depths, and used Monte Carlo sampling techniques assess the bounds and correlations between the estimated parameters. GPS velocities from this network constrain the total slip rate on the San Andreas fault system to be $39.6^{+1.5}_{-0.6}$ mm/yr (68.6% upper and lower confidence intervals). Slip rates on the individual faults are determined less precisely due to observed high correlations between estimated slip rates and locking depths, and between slip rates on adjacent faults. Our best fitting model estimates fault slip rates (all in mm/yr): San Andreas $17.4^{+2.5}_{-3.1}$, Ma'acama $13.9^{+4.1}_{-2.8}$, Bartlett Springs $8.2^{+2.1}_{-1.9}$, and locking depths (km): San Andreas fault $14.9^{+12.5}_{-7.1}$, Ma'acama fault $13.4^{+7.4}_{-4.8}$ except for shallow creep in the upper 5 km, and Bartlett Springs fault creeping at all depths. The estimated slip rate on the San Andreas fault is lower than all geologic estimates, although the 95% confidence interval overlaps the range of geologic estimates. The Ma'acama fault slip rate is greater than slip rate estimates for the Hayward or Rodgers Creek faults, its continuation to the south, from which we inferred a slip deficit large enough to generate a magnitude 7 earthquake today, posing a significant seismic hazard.

The study by *J. Murray et al*. (2014), updates these results using more recent data, a more complex fault geometry, and Bayesian inversion techniques to obtain *a posterior* probability density function for the full model space. They applied modest *a priori* constraints to the slip rates, and fixed the locking depths loosely based on seismicity. The estimated deep slip rates are similar to *Freymueller et al*. (1999), but have much smaller uncertainties (<1 mm/yr), due in part to the increased number of data and *a priori* slip-rate constraints, but more importantly due to the fixing of the locking depths, which eliminates some of the high correlations between the model parameters. Their Bayesian inversion approach provides a flexible method to assess the model uncertainties and to incorporate additional information, such as *a priori* constraints on locking depths based on seismicity, creep rates, etc.

GPS Data Processing

With the establishment of large continuously operating GPS networks over the last few decades, GPS time series and velocities are now available for over 13,000 stations around the globe. This explosion of observations has led to significant improvements in the ease and automation of GPS processing. Two different approaches to high-precision processing of GPS data were developed independently within the research community beginning in the 1980s, and both remain under active development today. The EarthScope PBO project compares and combines results from both techniques to provide high-quality data products for the research community interested in active deformation in North America.

The different approaches were initially motivated by the need to account for the drift of satellite and receiver clocks, which introduced significant biases in the estimated satellite-to-receiver ranges used to do precise positioning. One approach, adopted by GIPSY-OASIS developed at JPL, chose to explicitly estimate the clock drifts using Kalman filtering techniques. The use of the global reference network to estimate satellite orbit and clock information, plus refinements such as widelane-phase bias estimation (*Bertiger et al*., 2010), now enable the positions of stations to be estimated completely independently from other stations. This technique, generically known as Precise Point Positioning (PPP), enables rapid estimation of station positions and is well suited to handle the large data sets being generated by the growing continuous networks around the globe.

The alternative approach, adopted by GAMIT-GLOBK developed at MIT and Scripps, chose to double-difference observations between two satellites and two receivers, which eliminates the clock errors completely. However, forming the double differences requires relative observations between stations, which is usually performed on regionally clustered networks of stations where the short distances between

stations also helps to minimize differences due to atmospheric effects and improve resolution of phase biases. For computation efficiency of the least square inversions, the networks for an individual solution are usually be limited in size (<100 stations). However, the processing of the individual networks can be done in parallel and then combined into a single solution, making the processing of large numbers of stations tractable.

One issue that has been revealed by these efforts is the difference in structure of the position and velocity covariance estimates provided by the two techniques. Because nearby stations are subject to nearly the same orbital and atmospheric errors, we expect their position and velocity estimates to be spatially correlated. The network approach, which fully propagates measurement errors through the least sequares inversion, provides some measure of the intersite correlations. The PPP approach, which processes stations independently, essentially eliminates the intersite correlations. One goal of this project is to assess how important these differences are. Although many studies have shown how non-white noise (e.g., flicker noise and random walk) temporal correlations can affect the uncertainty estimates of station velocities (e.g., *Langbein and Johnson*, 1997; *Mao et al.,* 1999; *Williams et al.,* 2004; *Langbein*, 2008), relatively few studies have investigated the spatial correlations of the time series. Two studies (*Williams et al.,* 2004; *Amiri-Simkooei*, 2009) found high correlations (~0.8) between north, east, and vertical components at short distances that decreased to ~0.5 at 1000 km separation for global solutions, which is reasonable given that nearby sites are presumably affected similarly by orbital biases and atmospheric variations.

For many applications, having independently derived station velocities or displacements is adequate. However, strain-rate estimates are derived from gradients in the velocity field and therefore are much more sensitive to the intersite correlations. Consider a trivial example: finding the uncertainty of the difference between two variables x and y, with variance (square of the uncertainty) σ_x^2 and σ_y^2, and covariance σ_{xy}. Fully propagating the covariance of the variables shows that the variance of $x–y$ is $\sigma_{x\text{-}y}^2 = \sigma_x^2 - 2\sigma_{xy} + \sigma_y^2$. Ignoring the covariance σ_{xy} results in $\sigma_{x\text{-}y}^2$ equal to the sum of the variances of x and y. If the variables are positively correlated, the $\sigma_{x\text{-}y}^2$ is decreased, and vice versa. Strain-rate estimates are essentially derived from the same differencing operation (scaled by the distance between sites), and are therefore similarily affected by inter-site correlations. Because strain accumulation models are most sensitive to gradients in the velocity fields, we anticipate that the bounds that can be inferred for these models are also sensitive to the inter-site correlations.

For the PBO project, CWU processed the data from over 2000 stations using the GIPSY-OASIS PPP techniques, and NMT processed the data using the GAMIT-GLOBK network techniques. The results from both Analysis Centers are then compared and combined by Tom Herring at MIT, and provided to the research community by UNAVCO. The two techniques in general provide very similar results, and have proven valuable for identifying metadata and modeling problems, and for testing new modeling approaches. The methodology of the processing strategies, advantages afforded by comparing the two techniques, and preliminary results for both tectonic and non-tectonic deformation sources are described in detail by *Herring et al. (*2016). This multi-analysis center PBO/GAGE approach ended in September 2018 when NSF decided to fund only the CWU analysis during the next 5-year period. A final combination for daily station positions, time series, and velocities that is based on both CWU and NMT analyses was released in July 2019 (*see Products Section*). The products included in this release include daily positions in the solution-independent SINEX format for both the CWU and NMT analyses separately, as well as the combined PBO analysis. This release therefore provides a self-consistent, high-quality data set for examing the differences between the PPP and nework-based analysis approaches.

Strain accumulation estimation

One of the long-term goals of this project is to assess the effect of inter-station correlations on full three-dimensional finite-fault models representing the complex fault geometries in the San Francisco Bay area, and to fully incorporate survey-mode observations in the region by using both the USGS GIPSY results and analysis of the data using GAMIT/GLOBK. This goal proved to be beyond the scope of investigation during the project period given the significant computational burdens required to do a full reprocessing of the PBO/GAGE analysis of the entire 1996-2018 data set using updated satellite absolute phase center models and IGS14 orbital models, and a similar concurrent reanalysis performed by the USGS.

We therefore restricted our inter-station correlation assessment to a simple one-dimensional screw dislocation test case that allows for straightforward comparison of a limited number of model parameters.

The one-dimensional screw dislocation model, described by *Savage and Burford* (1973), assumes an infinitely long screw dislocation embedded in a homogeneous, elastic half-space. This formulation is equivalent to assuming that slip on a vertical fault is zero from the surface to a locking depth D, has a long-term slip rate s on the fault below D, which results in stations located x distance from the fault have fault-parallel surface velocities given by $v_i = (s/\pi)\ tan^{-1}(x_i/D)$. Estimation of both s and D is non-linear, but assume a locking depth D fixes the geometry, so that estimation of slip rate s from the observed station velocities is linear, enabling a simple weighted least-square inversion.

To approximate the infinite fault geometry, we consider a profile of stations spanning the creeping central San Andreas fault south of San Juan Bautista. The PBO/GAGE distributed velocity solutions do not provide the full covariance matrices for the station velocities, so we estimated the velocities using full-covariance daily station position solutions provided by CWU, NMT, and PBO. We chose the 10-yr time interval from 2008.169 to 2018.244, which corresponds to the period when the PBO network was fully operational and collecting data from 100-120 stations in the San Francisco Bay area (Figure 2).

Since the primary goal of the study is to assess inter-station correlations, we minimized the effect of temporal correlations by analyzing only a subset of the days. To determine a reasonable sampling interval, we tested different intervals and compared white noise versus colored noise daily scatter estimates, as determined by time series analysis tools provided by GAMIT/GLOBK. The color-noise estimate was derived from the FOGMEX first-order Gauss Markov estimate methodology described by *Floyd and Herring* (2020). The goal is to use an interval where the white noise estimates are roughly similar to the color-noise estimates. Tests of different sampling intervals suggested that after 32 days the red-noise component becomes minimal (Figure 3). Using 32-day intervals, we estimated station velocities based on 116 daily solutions in the 2008-2018 time frame.

We estimated the station velocities using the GLOBK analysis software following methods described in *Herring et al*. (2016). GLOBK uses a two-step Kalman filter approach. In the first step, daily position solutions with loosely constrained absolute positions, but tightly constrained relative positions, are combined to estimate loosely constrained position and velocity components. In the second step, the position and velocity reference frames are realized by allowing rotation and translation of these frames in order to minimize differences with a priori coordinates and velocities. This frame "stabilization" approach yields tightly constrained position and velocity solutions while providing great flexibility in their definition, such as allowing velocities to be estimated with respect to different tectonic plates. We estimated the velocities relative to the NAM14 reference frame developed for PBO/GAGE, which attempts to realize a fixed North America plate frame.

Use of GLOBK requires considerable care in defining a priori uncertainties and process noise for the Kalman filter. The PBO/GAGE daily solutions, made availble in SINEX format, include station positions but do not include other estimates parameters, such as atmospheric zenith delays, satellite orbits, earth-orientation parameters, etc. The positions are also provided in a tightly constrained fixed reference frame. To perform loosely constrained combinations, the station position covariance matrices are augmented with uncertainties that allow translations and rotations in the stabilization step. These augmented uncertainties are applied uniformly to all stations in the solution and essentially cancel out during the stabilization step, and have minimal effect on the inter-station correlations of interest in this study. However, the Kalman filter in GLOBK is particularly sensitive to numerical noise whose propagation during the Kalman filter steps can cause issues such as theoretically impossible negative chi-square increases when new data is added. In practice, this means that the number of stations included in the estimation must be limited to less than 80 for the current implementation of GLOBK. Estimation of velocities for large networks, such as for the 2000+ station PBO/GAGE solution, requires that the network be partitioned into smaller sub-networks for the loosely contrained combination of daily solutions. The resulting loosely contrained sub-network velocities can then be combined and stabilized to provide a final velocity solution for the entire network. One important implication of this approach is that the resulting covariance for the entire network has minimal inter-station correlations between the sub-networks, except as imposed by the common stations used to tie the sub-networks together. For the simple screw dislocation analysis, we limited the network to

the 30 stations shown in Figure 4, from which the 12 profile stations were chosen, to minimize the numerical precision errors and provide full velocity covariance estimation.

We estimated velocities and their covariances for the CWU, NMT, and combined PBO solutions. The velocity estimation accounted for offsets due to equipment changes, and earthquakes (such as the 2003 M6.6 San Simeon and M6.0 Parkfield events), with velocities assumed equal for each time series segment with no post-seismic transient deformation estimated. Horizontal velocities, shown for the PBO solution in Figure 4, had uncertainties ~0.1 mm/yr, similar to that found by the PBO/GAGE velocity product. The horizontal velocities for the 12 profile stations were then rotated into fault parallel (321 degree azimuth) and transverse components. The fault parellel componet velocities, plotted versus distance from the fault in Figure 5 have velocities ranging from ~14 mm/yr in the central valley to ~45 mm/yr at the coast relative to stable North America realized by the NAM14 reference frame. There is a sharp increase at the San Andreas fault implying a shallow locking depth on this creeping section of the fault.

To estimate the deep slip rate, we modified the screw dislocation model to include a reference frame velocity offset v_0 to account for the velocity of the position of San Andreas fault in the NAM14 frame, as well as fault origin offset x_0 to allow for uncertainties in the fault location, resulting in the model:

$$v_i = v_0 + (s/\pi)\, tan^{-1}((x_i + x_o)/D)$$

We performed a grid search to determine reasonable estimates on the non-linear parameters: $x_0 = -1.6$ km, and $D = 0.6$ km, indicating a very shallow locking depth, i.e., the fault in this region is almost entirely creeping at all depths. Fixing these parameters allows model parameters v_0 and s to be estimated using standard weighted least-square techniques: i.e., given d data, C_d=cov(d), and m model parameters related linearly to d by $d = Gm$, then $m = (G^T C_d^{-1} G)^{-1} G^T C_d^{-1} d$, and cov($m$) = $(G^T C_d^{-1} G)^{-1}$. To evaluate the fit, we determine the normalized root-mean-square (NRMS) of the residuals. The NRMS is the square root of the weighted residual chi-square divided by the degrees of freedom (12 data – 2 model parameters). The NRMS provides a measure of how well the observed scatter agrees with the formal data uncertainties. Ideally NRMS~1, while high values of NRMS may indicate blunders or under-estimated formal uncertainties.

The results for the estimates are shown in Table 1. All models return similar values for the velocity offset of 29.7 ± 0.02 mm/yr and a deep sliprate of 31.1 ± 0.04 mm/yr, which is consistent with deep sliprates estimated by more detailed finite-fault modeling (e.g., *Rolandone et al.,* 2008). The NMT solution is very similar to the PBO solution both in terms of parameter values and uncertainties. The CWU solution has larger uncertainties than the NMT solution, suggesting that the NMT solution has stronger weight in the combined PBO solution. To perform the combined PBO solution, the CWU and NMT solution uncertainties are individually scaled to make the scatter of their residuals consistent with their formal uncertainties, and there is an explicit scale weighting factor between the CWU and NMT to attempt to account for some of the model differences. So it is difficult to compare the levels of their formal uncertainties. The larger CWU uncertainties results in a smaller NRMS of the fits (~11 for PBO and NMT versus 7.5 for the CWU). These NRMS are significantly higher than expected, due in part to misfits at stations including several near the coast that could be influenced by the San Gregonio-Hosgri fault system. The high NRMS could also be a result of underestimated formal errors in the velocity solutions. More realistic uncertainties, at least more consistent with the residual scatter, can be obtained by multiplying the formal uncertainties by NRMS. In this case, the sliprate uncertainty is 0.4 mm/yr for all the solutions. Thus, in terms of providing an estimate for sliprate uncertainty assuming a simple screw dislocation model, all the solutions are indisinguishable.

To assess the effect of inter-station correlations independent of the uncertainty scaling issues between the different solutions, we examine the correlation matrices of the velocity solutions. Figure 6 shows colored images of the correlation matrices for the fault parallel velocity components of the 12 profile stations. The non-dimensional correlation matrices were derived from the velocity covariance matrices, where $corr_{i,j} = cov_{i,j}/sqrt(cov_{i,i}cov_{j,j})$. Correlation coefficients range from -1 (anti-correlated) to 1 (fully-correlated). For the 3 solutions, all inter-station correlations are less than 0.25 and greater than -0.15, so neither strongly correlated or anti-correlated. The NMT correlations are generally slightly more positive than the PBO or CWU correlations, particularly for stations west of the SAF, but the differences are less than 0.2. The CWU and PBO correlations are very similar. To see if these differences are significant, we used the NMT correlation matrix scaled by the PBO formal uncertainties to generate a hybrid PBO/NMTcor

covariance matrix. The estimated results using this hybrid covariance, and a similar test using the CWU correlation matrix, are given in Table1. The CWU hybrid results are very similar to the PBO results, as expected. The NMT hybrid results have slightly (10%) higher model uncertainties (0.0415 vs 0.0378 mm/yr for sliprate.

Conclusions

During the project period, we completed analysis of over 2000 GNSS stations in the western U.S. for the PBO/GAGE project, culminating with an official release of NMT, CWU, and combined PBO daily station positions, time series, and velocity products in June 2019. We have used the results of these analyses to perform an initial assessement of inter-station correlations using the a one-dimensional screw dislocation model applied to a linear array across the creeping section of the San Andreas fault in central California. This simple linear inversion model for fault slip allows for direct evaluation of the effect of the different correlation matrices. Our preliminary results suggest that the differences do not significantly affect realistically scaled uncertainties on the estimated deep sliprate. The estimated inters-station correlations of the fault-parallel velocity components are less than 0.25 and greater than -0.15, indicating a lack of strong correlations or anti-correlations between the stations. The network-derived NMT correlations have slightly more positive inter-correlations (0.05-0.2) than the PPP-derived CWU solutions, which results in a modestly higher (10%) sliprate uncertainty, but more tests are required to determine if this is significant.

Future work on this issue could focus on several topics. It is unclear whether the existing covariance structure of the network solutions provides a realistic characterisation of inter-station correlations. Many steps are required to obtain velocities which can potentially reduce the correlations, such as fixing orbit parameters and earth-orientation parameters. Large networks require division into subnetworks to avoid numerical instabilities, which results in loss of inter-station correlations between the subnetworks. Frame stabilization techniques, similar to regional filtering (e.g., *Wdowinski et al*., 1997), which the USGS currently includes in their analysis, or Helmert transformations to realize reference frames have been found to decrease the correlations to less than 0.5. Development of a realistic model of the spatial correlations

The simple screw dislocation model tested in this project provides a zeroth order assessment of the potential impact of spatial correlations. Although the results suggest they may not be significant, it would still be interesting to do the comparison on a the more realistic three-dimensional fault model where correlations in the model parameters, such as between parallel faults, might be more critical. Finally, inter-station correlations may be more significant for differential strain-rate (as opposed to individual station velocity) studies, such as analysis of short-baseline differential baselines crossing the San Andreas fault (e.g., PEA1-PEA2 in the creeping zone).

TABLE 1. Velocity offset (mm/yr), Deep sliprate (mm/yr), and NRMS fit for screw dislocation model

	Velocity offset	Vel. off. sigma	Sliprate	Sliprate sigma	NRMS
PBO	29.7440	0.0210	31.0715	0.0378	11.1823
NMT	29.7256	0.0196	31.0529	0.0375	11.1708
CWU	29.5978	0.0300	31.1154	0.0563	7.4824
PBO/NMTcor	29.7797	0.0213	31.0294	0.0415	11.0345
PBO/CWUcor	29.7242	0.0205	31.1117	0.0385	11.0613

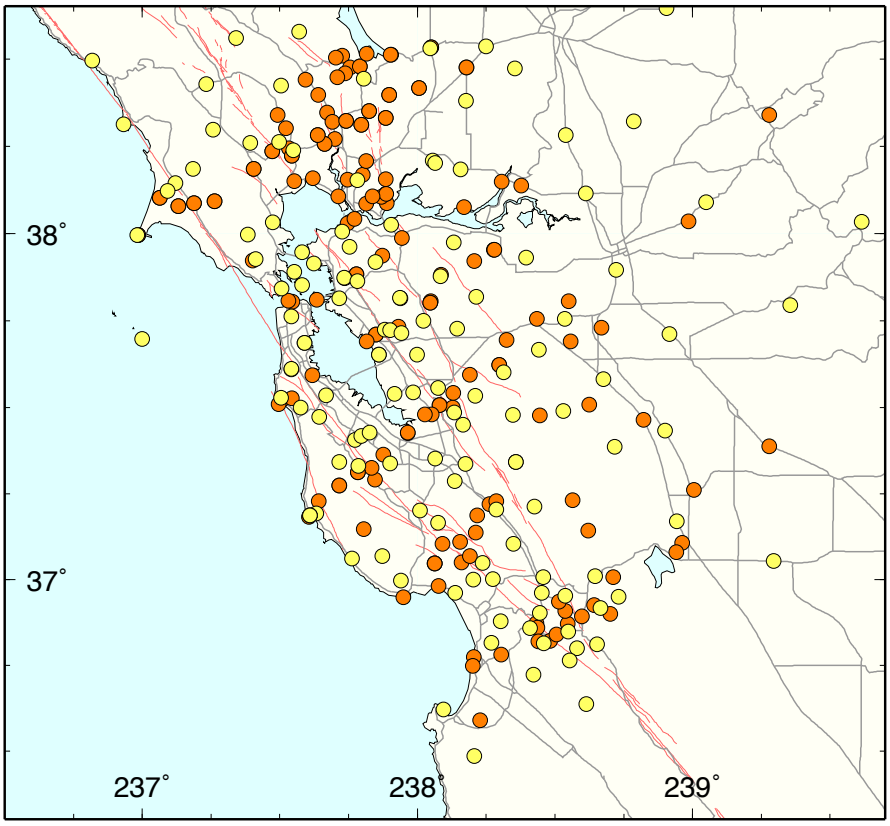

Figure 1. GPS sites (circles) in the San Francisco Bay area, including 130 continously operating GPS stations (yellow) and 290 survey-mode benchmark sites (orange) that are typically occupied for several days every few years. Active faults are shown by red lines and roads are shown in grey lines. The most NW sites are in the Sierra foothills. Many continous sites, such as in the Great Valley, were added during 2005-2008 as part of the EarthScope Plate Boundary Observatory project. This distribution of sites provides possible improvements to constraining strain accumulation of faults in the East Bay and Sacramento-San Joaquin Delta regions.

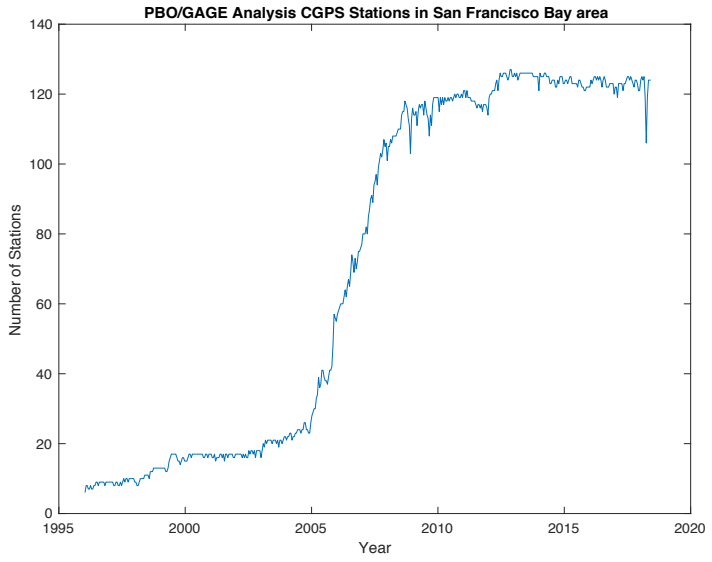

Figure 2. CGPS stations in the San Francisco Bay area included in the PBO/GAGE Repro2 analysis. Data from the ten-year period from 2008.2-2018.2, which continuously includes 105-127 stations, were used to estimate velocities and assess the effect of spatial covariance on strain accumulation estimations.

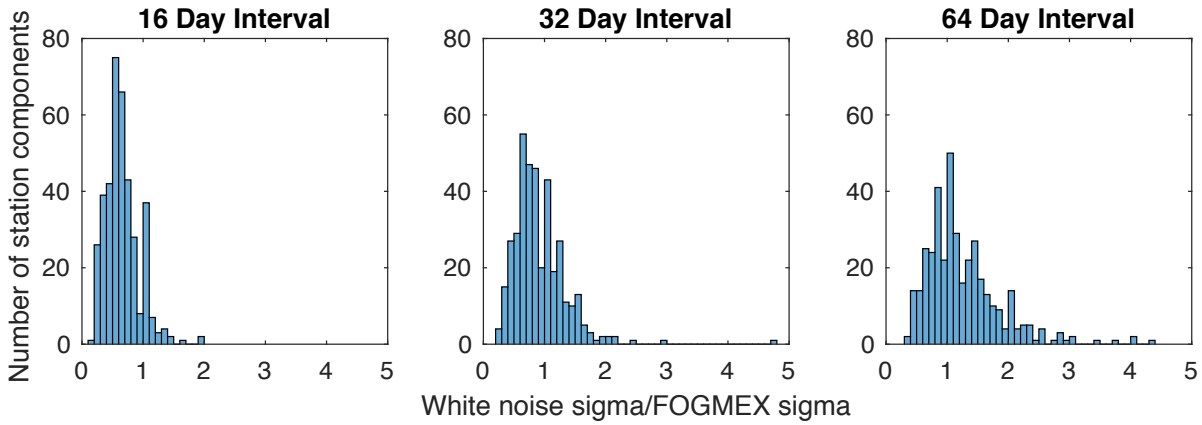

Figure 3. Comparison of white noise standard deviation (sigma) of component estimates relative to FOGMEX estimates (*Floyd and Herring*, 2020) for sampling intervals of 16, 32, and 64 days. 16-day estimates are generally low. 32 and 64-day estimates cluster around the preferred 1 value, but 64-day estimates have a broader tail towards higher value, probably due to the fewer data values used in the estimate. We used the 32-day interval in this study.

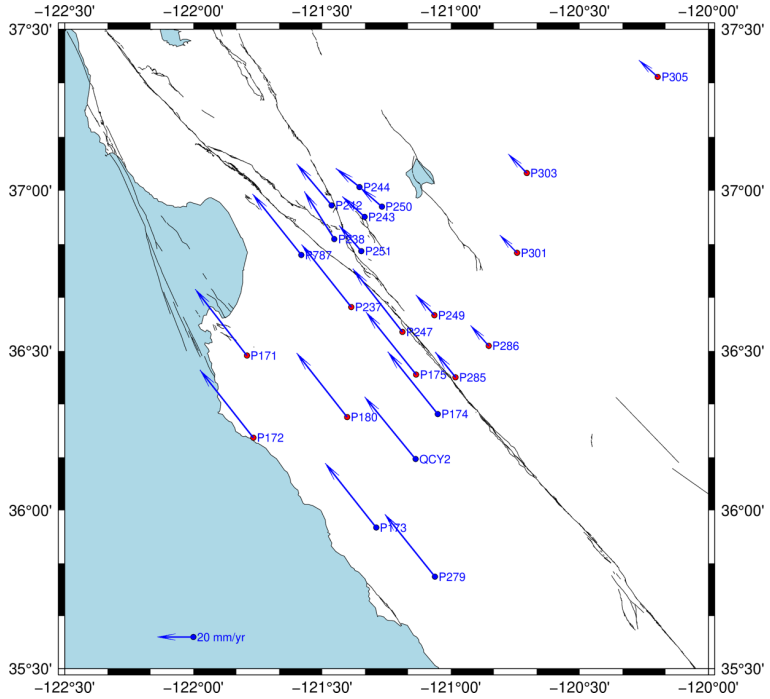

Figure 4. Horizontal velocities of GPS sites spanning the creeping section of the San Andreas fault estimated using the PBO combined position solutions from 2008-2018 relative to the NAM14 North America plate fixed reference frame. 95% confidence ellipses at tips of arrows indicate the low uncertainty values. Blue circles: stations included in the velocity frame realization. Red circles: profile stations used to estimate the deep sliprate and locking depth screw dislocation model.

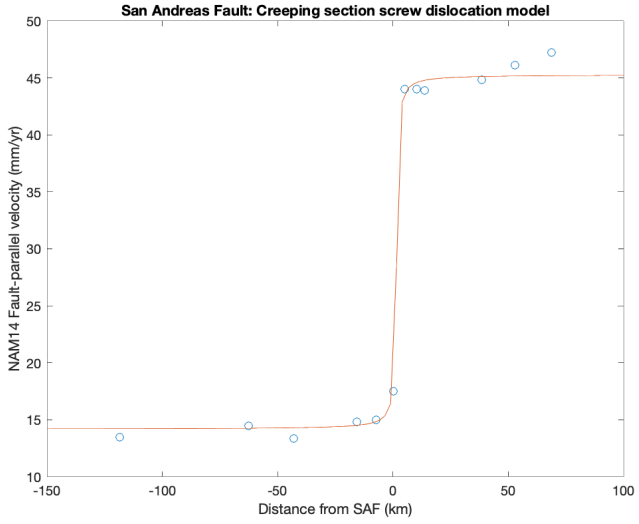

Figure 5. Fault-parallel velocities (circles) of GPS sites spanning the creeping section of the San Andreas fault used to estimate deep sliprate and locking depth screw dislocation model (red line). Formal errors are less than the radius of the circles. Distances are relative to a nominal fault location, with negative values towards the central valley and positive values towards the coast.

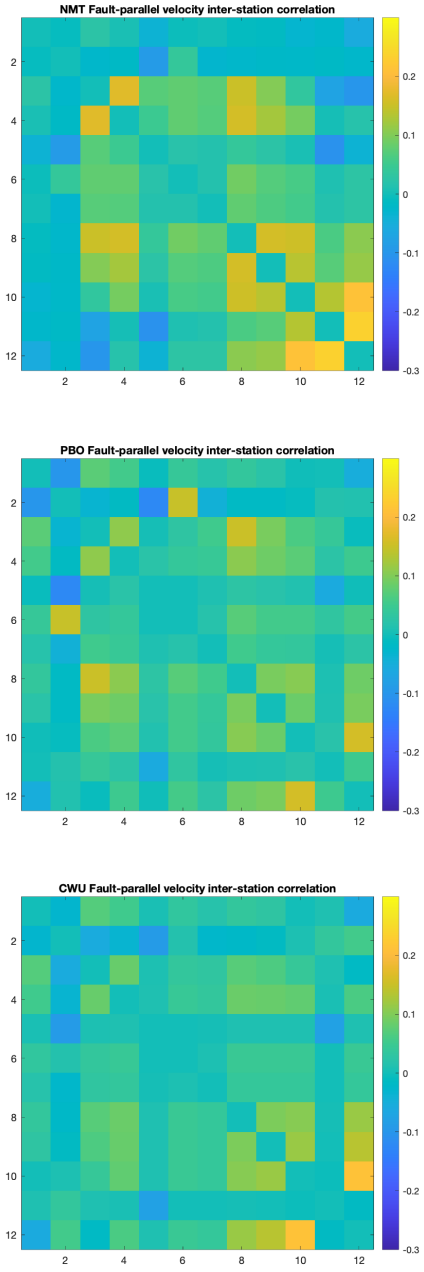

Figure 6. Correlation matrix values for the fault-parallel profile site velocities, ranging from -0.3 to 0.3 (colorbar) for the NMT (top), PBO (middle), and CWU (bottom) solutions. The diagonal value of 1.0 is set to 0.0 to highlight low inter-station correlations. Stations are ordered by distance from the San Andreas fault, with station 1 (P305) in the central valley, and station 12 (P172) on the coast. The CWU correlations, derived from PPP solutions, are similar to the PBO solution, while the NMT have higher inter-station correlations, particularly for sites southwest of the fault in the lower right corner of the matrix.

Acknowledgment and Disclaimer

This material is based upon work supported by the U.S. Geological Survey under Grant No. G17AP00005. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Geological Survey. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Geological Survey.

Project Data and Publications

No publications have resulted yet directly from funding provide by this grant, as significant resources and effort were required to reprocess the entire PBO/GAGE data set prior to release of station position, time series, and velocity products used in this project. The project data is derived from SINEX files for the PBO, NMT, and CWU (version repro NAM14) daily solutions, available at:

ftp://data-out.unavco.org/pub/products/sinex

These SINEX files were used to generate the final PBO/GAGE velocity field utilizing both CWU and NMT solutions. Description and ancillary files for the 2019 GAGE NAM14 Combined Velocity Field, released 2019-06-12, available at:

https://www.unavco.org/data/gps-gnss/derived-products/docs/GAGE_GNSS_Velocity_Field_Release_Notes_20190612.pdf
https://www.unavco.org/data/gps-gnss/derived-products/docs/processing-files/DOI_180915.zip

The analysis methodology used to process the PBO/GAGE velocity field is fully described in:

Herring, T. A., T. I. Melbourne, M. H. Murray, M. A. Floyd, W. M. Szeliga, R. W. King, D. A. Phillips, C. M. Puskas, M. Santillan, and L. Wang, Plate Boundary Observatory and related networks: GPS data analysis methods and geodetic products, *Rev. Geophys., 54*, doi:10.1002/2016RG000529, 2016.

References

Amiri-Simkooei, A. R., Noise in multivariate GPS position time-series. J. Geodesy, 83(2), 175-187, 2009.

Bertiger, W., Desai, S. D., Haines, B., Harvey, N., Moore, A. W., Owen, S., & Weiss, J. P. Single receiver phase ambiguity resolution with GPS data. Journal of Geodesy, 84(5), 327-337, 2010.

Bürgmann, R., P. Segall, M. Lisowski, and J. Svarc, Postseismic strain following the 1989 Loma Prieta earthquake from GPS and leveling measurements, J. Geophys. Res., 102(B3), 4933–4955, doi:10.1029/96JB03171, 1997.

Bürgmann, R., D. Schmidt, R. M. Nadeau, M. d'Alessio, E. Fielding, D. Manaker, T. V. McEvilly, and M. H. Murray, Earthquake potential along the northern Hayward fault, California, Science, 289, 1178-1182, 2000.

Chaussard, E., Bürgmann, R., Fattahi, H., Nadeau, R. M., Taira, T., Johnson, C. W., and Johanson, I.. Potential for larger earthquakes in the East San Francisco Bay Area due to the direct connection between the Hayward and Calaveras Faults. Geophys. Res. Lett., in press, 2015.

D'Alessio, M. A., I. A. Johanson, R. Bürgmann, D. A. Schmidt, and M. H. Murray, Slicing up the San Francisco Bay area: Block kinematics from GPS-derived surface velocities, J. Geophys. Res., 110, B06403, doi:10.1029/2004JB003496, 2005.

DeMets, C., R. G. Gordon, D. F. Argus, and S. Stein, Effects of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions, Geophys. Res. Lett., 21, 2191-2194, 1994.

Dengler, L., Moley, K., McPherson, R., Pasyanos, M., Dewey, J., and Murray, M., The September 1, 1994 Mendocino Fault earthquake, Calif. Geol., 48, 43–53, 1995.

Dickinson, W. R. and W. S. Snyder, Geometry of triple junctions related to San Andreas transform, J. Geophys. Res., 84, 561-572, 1979.

Evans, E. L., J. P. Loveless, and B. J. Meade, Geodetic constraints on San Francisco Bay Area fault slip rates and potential seismogenic asperities on the partially coupled Hayward fault, Journal of Geophysical Research, 117, B03410, doi:10.1029/2011JB008398, 2012.

Field, Edward H., et al. Long-term time-dependent probabilities for the third Uniform California Earthquake Rupture Forecast (UCERF3). BSSA 105.2A, 511-543, 2015.

Floyd, Michael A., and Thomas A. Herring, Fast Statistical Approaches to Geodetic Time Series Analysis, In Geodetic Time Series Analysis in Earth Sciences, pp. 157-183. Springer, Cham, 2020.

Freymueller, J. T., M. H. Murray, and P. Segall, Kinematics of the Pacific-North America plate boundary zone, northern California, J. Geophys. Res, 104, 7419-7442, 1999.

Furlong, K. P., Thermal-rheologic evolution of the upper mantle and the development of the San Andreas fault system, Tectonophysics, 223, 149-164, 1993.

Furlong, K. P., and R. Govers, Ephemeral crustal thickening at a triple junction: The Mendocino crustal conveyor, Geology, 27, 127-130, 1999.

Furlong, K. P., W. D. Hugo, and G. Zandt, Geometry and evolution of the San Andreas fault zone in northern California, J. Geophys. Res., 94, 3100-3110, 1989.

Herring, T., M. Floyd, R. King, T. Melbourne, W. Szeliga, M. Murray, D. Phillips, C. Puskas, F. Boler, C. Meertens, G. Mattioli, GPS data analysis and results from the Geodesy Advancing Geosciences and EarthScope (GAGE) project , Abstract G13A-0511 presented at 2014 Fall Meeting, AGU, San Francisco, Calif., 15-19 Dec., 2014.

Herring, T. A., T. I. Melbourne, M. H. Murray, M. A. Floyd, W. M. Szeliga, R. W. King, D. A. Phillips, C. M. Puskas, M. Santillan, and L. Wang, Plate Boundary Observatory and related networks: GPS data analysis methods and geodetic products, Rev. Geophys., 54, doi:10.1002/2016RG000529, 2016.

Jachens, R. C. and A. Griscom, Three-dimensional geometry of the Gorda plate beneath northern California, J. Geophys. Res., 88, 9375-9392, 1983.

Johanson, I. A., and R. Bürgmann, Creep and quakes on the northern transition zone of the San Andreas fault from GPS and InSAR data, Geophys. Res. Lett., 32, L14306, doi:10.1029/2005GL023150, 2005.

Langbein, J., Noise in two-color electronic distance meter measurements revisited, J. Geophys. Res., 109, B04406, doi:10.1029/2003JB002819, 2004.

Langbein, J., Noise in GPS displacement measurements from Southern California and Southern Nevada, J. Geophys. Res. , 113(B5), B05405, 2008.

Langbein, J., and H. Johnson, Correlated errors in geodetic time series: Implications for time-dependent deformation. J. Geophys. Res., 102(B1), 591-603, 1997.

Lisowski, M., J.C. Savage, and W.H. Prescott, The velocity field along the San Andreas fault in central and southern California, J. Geophys. Res., 96, 8369-8389, 1991.

Manaker, D. M., R. Bürgmann, W. H. Prescott, and J. Langbein, Distribution of interseismic slip rates and the potential for significant earthquakes on the Calaveras fault, central California, J. Geophys. Res., 108(B6), 2287, doi:10.1029/2002JB001749, 2003.

Mao, A., C. G. Harrison, and T. H. Dixon, Noise in GPS coordinate time series. J. Geophys. Res., 104(B2), 2797-2816, 1999.

Matthews, M. V., and P. Segall, Statistical inversion of crustal deformation data and estimation of the depth distribution of slip in the 1906 earthquake, J. Geophys. Res., 98, 12,153--12,163, 1993.

Matsu'ura, M., D. D. Jackson, and A. Cheng, Dislocation model for aseismic crustal deformation at Hollister, California, J. Geophys. Res., 91, 2661--2674, 1986.

McCaffrey, R., Crustal block rotations and plate coupling, in Plate Boundary Zones, S. Stein and J. T. Freymueller (eds.), AGU Geophysical Monograph30, 101-122, 2002.

McClusky, S., et al., Global Positioning System constraints on plate kinematics in the eastern Mediterranean and Caucasus, J. Geophys. Res., 105, 5695--5719, 2000.

Meade, B. J., and J. P. Loveless, Block modeling with connected fault-network geometries and a linear elastic coupling estimator in spher- ical coordinates, Bull. Seismol. Soc. Am., 99(6), 3124–3139, doi:10.1785/0120090088, 2009.

Meade, B. J., B. H. Hager, S. C. McClusky, R. E. Reilinger, S. Ergintav, S., O. Lenk, A. Barka, and H. Ozener, Estimates of Seismic Potential in the Marmara Sea Region from Block Models of Secular Deformation Constrained by Global Positioning System Measurements, Bull. .Seismol. Soc. .Am., 92, 208-215, 2002.

Murray, J. R., S. E. Minson, and J. L. Svarc, Slip rates and spatially variable creep on faults of the northern San Andreas system inferred through Bayesian inversion of Global Positioning System data, J. Geophys. Res. Solid Earth, 119, 6023–6047, doi:10.1002/2014JB010966, 2014.

Murray, M.H., and M. Lisowski, Crustal deformation near the Mendocino Triple Junction, EOS Trans. AGU, 76(46), Fall Meeting Suppl., 627, 1995.

Murray, M. H., and P. Segall, Modeling broadscale deformation in northern California and Nevada from plate motions and elastic strain accumulation, Geophys. Res. Lett., 28, 4315-4318, 2001.

Murray, M. H., Marshall, G. A., Lisowski, M., and Stein, R. S., The 1992 M=7 Cape Mendocino, California, earthquake: Coseismic deformation at the south end of the Cascadia megathrust, J. Geophys. Res., 17,707-17,725, 1996.

Niemi, T. M., and N. T. Hall, Late Holocene slip rate and recurrence of great earthquakes on the San Andreas fault in northern California, Geology, 20, 195-198, 1992.

Prentice, C. S., Earthquake geology of the northern San Andreas fault near Point Arena, California, Ph. D. thesis, 349 pp., Calif. Inst. of Technology, Pasadena, 1989.

Prentice, C. S., W. H. Prescott, R. Langridge, and T. Dawson, New geologic and geodetic slip rate estimates on the North Coast San Andreas fault: Approaching agreement?, Seismol. Res. Lett., 72, 282, 2001.

Prescott, W. H., J. C. Savage, J. Svarc, and D. Manaker, Deformation across the Pacific-North America plate boundary near San Francisco, California, J. Geophys. Res., 6673-6682, 2001.

Rolandone, F., R. Bürgmann, D. C. Agnew, I. A. Johanson, D. C. Templeton, M. A. d'Alessio, S. J. Titus, C. DeMets, and B. Tikoff, Aseismic slip and fault-normal strain along the central creeping section of the San Andreas fault, Geophys. Res. Lett. 35, 2008.

Savage, J. C., A dislocation model of strain accumulation and release at a subduction zone, J. Geophys. Res., 88, 4984--4996, 1983.

Savage, J. C., and R. O. Burford, Geodetic determination of relative plate motion in central California, J. Geophys. Res., 78, 832--845, 1973.

Schmidt, D. A., R. Bürgmann, R. M. Nadeau, and M. d'Alessio, Distribution of aseismic slip rate on the Hayward fault inferred from seismic and geodetic data, J. Geophys. Res., 110, B08406, doi:10.1029/2004JB003397, 2005.

Segall, P., Integrating geologic and geodetic estimates of slip rate on the San Andreas fault system, Int. Geol. Rev., 44, 62-82, 2002.

Thatcher, W., Strain accumulation and release mechanism of the 1906 San Francisco earthquake, J. Geophys. Res., 80, 4862-4872, 1975.

Thatcher, W., G. Marshall, and M. Lisowski, Resolution of fault slip along the 470 km long rupture of the great 1906 earthquake and its implications, J. Geophys. Res., 102, 5353-5367, 1997.

Wdowinski, S., Y. Bock, J. Zhang, P. Fang, and J. Genrich, Southern California permanent GPS geodetic array: Spatial filtering of daily positions for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake, J. Geophys. Res., 102(B8), 18,057–18,070, doi:10.1029/97JB01378, 1997.

Williams, S. D. P., Y. Bock, P. Fang, P. Jamason, R. M. Nikolaidis, L. Prawirodirdjo, M. Miller, and D. J. Johnson, Error analysis of continuous GPS position time series, J. Geophys. Res., 109(B3), B03412, 2004.

WGNEP (Working Group on Northern California Earthquake Potential), Database of potential sources for earthquakes larger than magnitude 6 in northern California, U. S. Geological Survey Open File Report 96-705, 1996.

WG99, Earthquake probabilities in the San Francisco Bay region: 2000 to 2030:A summary of findings, USGS Open File Rpt., 99-417, 1999.

Yagi, Y., and Y. Fukahata, Introduction of uncertainty of Green's function into waveform inversion for seismic source processes. Geophys. J. Int., 186(2), 711-720, 2011.